Theoretical Computer Science 129 (1994) 369–383
Elsevier

Tiling and local rank properties of the Morse sequence

S. Ferenczi
LMD, CNRS-UPR 9016, Case 930, 163 Avenue de Luminy, F-13288 Marseille, France

Communicated by D. Perrin
Received December 1990
Revised February 1991

Abstract

Ferenczi, S., Tiling and local rank properties of the Morse sequence, Theoretical Computer Science 129 (1994) 369–383.

We study some combinatoric properties of the Morse sequence, linked with its ergodic properties of local rank one and local funny rank one; we show that the maximum part of the Morse sequence that may be covered by disjoint translates of one word is exactly of density 2/3, even allowing for some errors in the tiling; when we replace words by patterns (words with holes), 2/3 can be replaced by at least 5/6.

In this paper, let σ be the substitution given by

$$\sigma(0)=01, \qquad \sigma(1)=10$$

and $u=(u_n,\ n\in\mathbb{N})$ the fixed point of σ beginning by 0, that is the classical Morse sequence:

$$0110100110010110100101100110100 1\cdots$$

First it was introduced by Prouhet [10], then independently by Thue [12] and Morse [8]. Its combinatoric properties were extensively studied, the reader will find in [7] a survey of these and a complete bibliography; but the Morse sequence may be viewed also as a *measure-preserving dynamical system*: namely (Ω, T, μ), where T is the shift, $(Tu)_n = u_{n+1}$; Ω is the closed orbit of u under T; μ is the unique ergodic measure

Correspondence to: S. Ferenczi, LMD, CNRS-UPR 9016, Case 930, 163 Avenue de Luminy, F-13288 Marseille, France.

0304-3975/94/$07.00
SSDI 0304-3975(93)E0181-3

preserving T. This system, along with other systems defined by substitutions, has been extensively studied, see for example [11].

The system associated with the Morse sequence has interesting ergodic properties: it has *simple spectrum in* $\mathscr{L}^2$, which means that the associated operator on $\mathscr{L}^2(X)$ defined by $Uf=f\circ T$ satisfies the following property: there exists an element f of $\mathscr{L}^2(X)$ such that $\mathscr{L}^2(X)$ is the closed linear subspace generated by the $U^n f$, $n\in\mathbb{Z}$. It also belongs to the well-known category of *finite rank systems* [9], see Definition 2.

It was an important step in the chain of ergodic examples answering the question "which kind of ergodic systems have simple spectrum?"; the first conjecture was that simple spectrum was equivalent to *rank one* (see Definition 2) and Del Junco [2] showed that the Morse sequence gave a counter-example; the notion of rank one was then weakened to *local rank one* ([3], see Definition 4) and *funny rank one* ([4], see Definition 11), and the standing question is now whether simple spectrum is equivalent to funny rank one. In this paper, we investigate the properties of the Morse sequence in regard to these notions, as it is a likely candidate for a counterexample to this last question; this has led us to consider some combinatorial properties connected to the *tiling* of the Morse sequence by its factors, this is of *covering it by disjoint translates of finite objects.*

Namely, considering the local rank one properties of the Morse sequence led us to compute which proportion of the sequence can be tiled by one finite *word* (see Definition 1); this turns out to be exactly 2/3, and is not increased if we allow small errors in the tiling; the words we use to get the best tiling come very naturally as the *n-words* (see Definition 1). To consider the funny rank one properties of the Morse sequence leads to the same questions with *patterns* (words with holes, see Definition 11) instead of words; we prove that patterns behave quite differently from words: one pattern can tile at least 5/6 of the Morse sequence; the well-known lemma about the nonoccurrence of $uVuVu$ for words V does not hold for patterns; the particular patterns we use to get the best tilings were, as far as we know, never studied before. The question of tiling the Morse sequence with patterns while allowing some errors is still completely open.

Definition 1. We call *word* any finite string $v_1v_2\ldots v_k$ of zeros and ones; we denote by VW the concatenation of the words V and W.

Between two words $V=v_1\ldots v_k$ and $W=w_1\ldots w_k$, we consider the *Hamming distance*

$$\bar{d}(V,W)=\frac{1}{k}\#(l \text{ such that } v_l\neq w_l)$$

(if we give properties of $\bar{d}(V,W)$, it will always, by definition, imply that the two words have the same length).

If v is a letter, we shall always note $v'=0$ if $v=1$, $v'=1$ if $v=0$; for $W=uvw\ldots$, we denote by W' the word $u'v'w'\ldots$

We call *n-words* the words $\sigma^n 0$ and $\sigma^n 1$, $n\geqslant 0$.

It is very trivial to note that the Morse sequence can be tiled by two words, 0 and 1, and this is not interesting from the ergodic point of view; hardly less trivial but already worth noting is the fact that u can be tiled by two words *of arbitrarily big length*, namely $\sigma^n 0$ and $\sigma^n 1$, which have length 2^n. This implies already that the dynamical system is of *rank* (at most) *two*.

We need to recall some definitions: a dynamical system of finite entropy can always, by the Krieger theorem, be equipped with a generating partition and so be viewed as a shift on sequences of a finite alphabet.

Definition 2. In this paper, we say that (Ω, T, μ) is of *rank at most* r if for all strictly positive ε, for every n, there exist r words W_i of length $l(W_i) \geq n$, and a finite collection of words W'_j satisfying

$$\text{for each } j, \quad \bar{d}(W_i, W'_j) < \varepsilon \text{ for at least one } i$$

such that, for all N big enough, with probability at least $1-\varepsilon$, the subsequences of length N of the elements of Ω are of the form

$$\varepsilon_1 W'_{j_1} \varepsilon_2 W'_{j_2} \dots \varepsilon_k W'_{j_k} \varepsilon_{k+1}$$

with $l(\varepsilon_1) + \dots + l(\varepsilon_{k+1}) < \varepsilon N$; a system is of rank r if it is of rank at most r and not of rank at most $r-1$.

This definition was derived by Del Junco [2] from an earlier, more geometrical, definition by Chacon [1]. Chacon proved also that a system of rank r has a spectral multiplicity bounded by r: $\mathscr{L}^2(X)$ is the closed linear subspace generated by the $U^n f_i$, $n \in \mathbb{Z}$, for at most r elements f_i of $\mathscr{L}^2(X)$. As we already mentioned, rank one implies simple spectrum.

We now come back to the Morse sequence. The paper of Del Junco proved that it is not of rank one: that is, it cannot be tiled by one word, even allowing for a small proportion of errors. However, it has simple spectrum, which means that it must possess some stronger property than rank two. Hence, we have been induced to compute the precise proportion of u which can be tiled by one word, with or without errors.

Definition 3. The word $W = v_1 v_2 \dots v_k$ is said to *occur* at place i in a (finite or infinite) string of letters $w_1 w_2 \dots$ iff $w_i = v_1, \dots, w_{i+k-1} = v_k$.

For a word V, we can define a *frequency* by $f(V) = \lim_{n \to +\infty} 2^{-n-1}$ (number of occurrences of V in $\sigma^n 0$ + number of occurrences of V in $\sigma^n 1$), and a *tiling frequency* by $\phi(V) = \lim_{n \to +\infty} 2^{-n-1}$ (maximum number of *disjoint* occurrences of V in $\sigma^n 0$ + maximum number of *disjoint* occurrences of V in $\sigma^n 1$). Both these limits make sense, as the sequences we consider are bounded from above and increasing.

If M is an infinite subset of $\mathbb{N}$, we call $E = (u_n, n \in M)$ an *infinite subset of the sequence* u, and we say E has *measure* $\mu(E) = m$ if M is of density m.

For a word V, the number $t(V)=\phi(V)l(V)$ is also the measure of the biggest subset of the Morse sequence u which can be tiled by V; we call it the *tiling* of the word V.

Definition 4. $F=\sup(a$ such that, for every n, there exists a word W of length $l(W)$ bigger than n, such that $t(W)\geqslant a)$,

$F^*=\sup(a$ such that, for all strictly positive ε, for every n, there exists a word W of length $l(W)\geqslant n$, and a finite collection of words W_i, such that $\bar{d}(W_i, W)<\varepsilon$ and there exists a subset of the Morse sequence of measure at least a which can be covered by disjoint occurrences of the $W_i)$.

Of course, the number F^* is the one that has ergodic interpretation, that is the one which is an invariant of metrical isomorphism; a system for which F^* is strictly positive is said to be of *local rank one*. These systems were studied in [3]; also, a well known and long unpublished result [6] attributed to Katok says that as soon as F^* is strictly greater than 1/2, the spectrum is simple.

We are now interested in computing F and F^* for the Morse sequence; note that F is closely connected with $\sup_{V \text{word}} l(V)\phi(V)$ (it actually equals this quantity provided the supremum is reached for arbitrarily long words); as f is a priori easier to compute than ϕ, we break our result into three steps.

Proposition 5. *For the Morse sequence,*

$$\sup_{V \text{ word}} l(V)f(V)=2/3.$$

Proof. Fix $n>0$ and take M bigger than n. The word $\sigma^n 0$ occurs in the word $\sigma^M e$, e being 0 or 1, at place j, if and only if j is odd and $\sigma^{n-1}0$ occurs in $\sigma^{M-1}e$ at place $(j+1)/2$, or, going back to the first stage, if and only if $j=k2^{n-1}+1$ and $\sigma^1 0=01$ occurs in $\sigma^{M-n+1}e$ at place k.

Hence, $f(\sigma^n 0)\geqslant 2^{-n+1}f(01)$. $f(01)$ has already been computed [2] but the same method gives it easily: 01 occurs in $\sigma^p e$ at place j iff

- $j=2k+1$ and 0 occurs at place k in $\sigma^{p-1}e$, or
- $j=2k$ and 11 occurs at place k in $\sigma^{p-1}e$;

11 occurs at place j in $\sigma^p e$ iff $j=2k$ and 01 occurs in $\sigma^{p-1}e$ at place k.

Hence, $f(01)=\frac{1}{2}(f(0)+f(11))$ and $f(11)=\frac{1}{2}(f(01))$; as $f(0)=\frac{1}{2}$, we get $f(01)=\frac{1}{3}$.

So we get $f(\sigma^n 0)\geqslant 1/(3.2^{n-1})$ for $n\geqslant 1$; as $\sigma^n 0$ is of length 2^n, we have $l(\sigma^n 0)f(\sigma^n 0)=\frac{2}{3}$ and the minoration side of our result.

Now, let W be a word of length $l(W)$ at least equal to 5; we shall compute $f(W)$.

Take m such that 2^m is much bigger than n. If W occurs in $\sigma^m e$ at place j, we partition $\sigma^m e$ in 1-words, $\sigma^1 e$. As the longest permitted words which do not contain 11 or 00 are 1010 and 0101, W must contain the word 11 or the word 00, which can only occur at the junction between two 1-words.

So j has only one possible parity, for example $j=2k+1$, and $W=|w_1w'_1|w_2w'_2|w_3w'_3|\ldots$ (the bars denoting the separations between 1-words), which means $U=w_1w_2w_3\ldots$ must occur at place k in $\sigma^{m-1}e$.

The same happens if $j=2k$, with $W=w_1|w_2w'_2|w_3w'_3\ldots$ and $U=w'_1w_2w_3\ldots$

So we have, as long as $l(V)$ is at least 5, $l(V)f(V)\leqslant l(U)f(U)$, where $\sigma(U)=pVs$, p being a strict prefix of 01 or 10, s being a strict suffix of 01 or 10, and U being strictly shorter than V. So we can go down like that to $l(U)\leqslant 4$. It is easy to compute that all permitted words of length 3 have frequency $\frac{1}{6}$, while words of length 4 have frequency smaller than $\frac{1}{6}$.

So we have proved that for any word W, $l(W)f(W)<\frac{2}{3}$, and the other side of the result. (We could also use directly the well-known fact that every word of length bigger than 4 is *uniquely interpretable*.) $\square$

Proposition 6. *For the Morse sequence,* $F=\frac{2}{3}$.

Proof. F is certainly smaller than the quantity in Proposition 5; on the other side, the supremum $\sup_{V\text{ word}} l(V)f(V)$ is reached for the words $\sigma^n 0$, $n\in\mathbb{N}$, and it is elementary to check that the occurrences of each of these words are disjoint, that is $f(\sigma^n 0)=\phi(\sigma^n 0)$; as these words have arbitrarily big length, we have $\frac{2}{3}=\sup_{V\text{ word}} l(V)\phi(V)=F$. $\square$

However, this can be deduced also from a more general result which has been known since [12]; for sake of completeness, we state this result and give a short proof slightly adapted from Hedlund and Morse [5].

Lemma 7. *For every word V in the Morse sequence, any two different occurrences of V are disjoint (hence $f(V)=\phi(V)$).*

Proof. If there are two nondisjoint occurrences of $V, V=AB=BC, A, B, C$ being nonempty words; it is easy to see that if W is the shortest of the words A and B and w its first letter, then the word WWw occurs in the Morse sequence; we shall show that a word of the form $wUwUw$, for any word U and letter w, cannot occur.

Suppose now that $wUwUw$ occurs and cut it into 1-words; taking into account the parity of $l(U)$, there are four possible ways to put separations: $|wU|wU|w$, $|wUw|U|w$, $w|U|wUw|$ and $w|Uw|Uw|$. The second and third suppose that U and wUw are unions of 2-words; this is false for $U=$ the empty word, and, if this is true for U nonempty, then U must be equal to $w'U'w'$ and U' and $w'U'w'$ are unions of 1-words, with $l(U')<l(U)$; going down to length zero, we see that this is impossible.

Hence, the first or fourth possibility must happen, hence either $U=w'U'$ and $|ww'U'|ww'U'|ww'|$ must occur (with this location of the separations) or else $U=U'w'$ and $|w'wU'|w'wU'|w'w|$ must occur; hence, $U'=\sigma(U'')$, and some $vU''vU''v$ must occur for some letter v and word U'' with $l(U'')<l(U)$; it just remains to check that this property is false for U empty, that is 000 and 111 do not occur. $\square$

To compute F^*, we need the following result from [2], for which we give Del Junco's proof.

Lemma 8. *If A and B are two n-words, and if C is a word occurring at place i in an M-word for an $M>n$ and satisfying $\bar{d}(AB,C)\leqslant\frac{1}{8}$, then*

$$C=AB \text{ and } i \text{ is a multiple of } 2^n.$$

(There are no neighbours in the Morse sequence.)

Proof. The result is empty for 0-words. Suppose it is true for $(n-1)$-words and let A, B, C, i be as in the hypothesis.

Then $C=C'C''$, with $\bar{d}(A,C')$ or $\bar{d}(B,C'')$ smaller than $\frac{1}{8}$; as both A and B are made of two $(n-1)$-words, the recursion hypothesis implies already that $i=k.2^{n-1}$; so C' and C'' are $(n-1)$-words, and are at a distance $\bar{d}<\frac{1}{2}$ of the corresponding $(n-1)$-words of A and B; as the distance $\bar{d}$ between $(n-1)$-words is 0 or 1, we must have $C=AB$.

It remains to prove that k is even; but if it is odd, if we cut A, B, and C in $(n-1)$-words, we get $C=UVV'W$, $A=XX'$, $B=YY'$, and $U'UVV'WW'$ must occur; hence $U'=V=W$, and $VV'VV'VV'$ occurs, hence 000 or 111 must occur, which is false. $\square$

Proposition 9. *For the Morse sequence, $F^*=\frac{2}{3}$.*

Proof. We use the techniques of [2]; let $a=F^*$. We take ε, n, W, the W_i, $1\leqslant i\leqslant p$, as in the definition; let $N=l(W)$.

We fix an integer b, and a δ smaller than $2^{-b}/100$; we choose an integer n much bigger than 2^b and apply the definition of F^* for δ and n; this gives a word W of length $l(W)=N$; we choose integers r and k such that $k2^r\leqslant N<(k+1)2^r$ and $2^b\leqslant k\leqslant 2^{b+1}-1$.

In W we see $k-1$ occurrences of words $\sigma^r e$, let $U_1, U_2, \ldots, U_{k-1}$ be these words. Now, if W is at a distance $\bar{d}<\delta$ of some word occurring in some $\sigma^p e$, then each subword U_iU_{i+1} is at a distance $\bar{d}<\frac{1}{8}$ of a word occurring in $\sigma^p e$.

Then, by Lemma 8, each U_iU_{i+1} must occur *exactly* in $\sigma^p e$. Which means each W_i must contain the word $U_1U_2\ldots U_{k-1}$ and, returning to the definition, we must have

$$a-\delta\leqslant Nf(U_1U_2\ldots U_{k-1}).$$

However, because this word has length $(k-1)2^r$, we use the previous paragraph to write

$$a-\delta\leqslant\frac{2N}{3(k-1)2^r}\leqslant\frac{2(k+1)}{3(k-1)}\leqslant\frac{2(2^b+1)}{3(2^b-1)}.$$

b and δ being arbitrary, we conclude that $a\leqslant\frac{2}{3}$, which gives the value of F^*, as of course F^* is bigger than F. $\square$

Corollary 10. *The associated dynamical system has simple spectrum.*

(This was known from [2], but by totally different methods, using spectral techniques and the expression of the dynamical system (X, T, μ) as a finite extension of a dyadic translation.)

So the problem of tiling the Morse sequence with one word is completely solved. Now, instead of considering words, we could consider "word with holes"; for example, instead of using the word 0110, we could try to tile u by the *pattern* (0, 1, six spaces, 1, 0) which we shall denote by $01 \overset{6}{\ldots} 10$; we shall see that this improves the tiling possibilities while keeping an ergodic significance.

Definition 11. A *pattern*, noted $v_1 \overset{a_1}{\cdots} v_2 \overset{a_2}{\cdots} \cdots \overset{a_{k-1}}{\cdots} v_k$ is a finite string of zeros, ones, and spaces. In this notation, meaning one digit v_1 followed by a_1 spaces then one digit v_2, etc., we allow some of the a_i to be equal to zero (if they are all equal to zero, the pattern is a word). We say k is the *weight* of the pattern, denoted by $p(V)$, and $(k + a_1 + \cdots + a_{k-1})$ is the *length* $l(V)$ of the pattern; by convention, we ask that a pattern always ends by a 0 or 1, so that the length is not ambiguous.

The pattern $v_1 \overset{a_1}{\cdots} v_2 \overset{a_2}{\cdots} \cdots \overset{a_{k-1}}{\cdots} v_k$ is said to *occur* at place i in the string $w_1 w_2 \ldots$ iff $w_i = v_1$, $w_{i+a_1+1} = v_2$, $w_{i+a_1+a_2+2} = v_3$, etc.

Two patterns V and V' of length l can be considered as words of length l on the alphabet (0, 1, "space"); we define $\bar{d}(V, V')$ as the Hamming distance between these two words.

We can now define, with the same definitions as for words, a *frequency* $f(V)$ and a *tiling frequency* $\phi(V)$ for any pattern V; we define also a *tiling* $t(V)$ as $\phi(V)p(V)$, or equivalently as the maximum proportion of the Morse sequence which can be tiled by V.

Then we can define

- $T = \sup(a$ such that, for every n, there exists a pattern V of weight $p(V)$ bigger than n, such that $t(V) \geq a)$,
- $T^* = \sup(a$ such that, for all strictly positive ε, for every n, there exists a pattern V of weight $p(V) \geq n$, and a finite collection of patterns V_i, such that $\bar{d}(V_i, V) < \varepsilon$ and there exists a subset of the Morse sequence of measure at least a which can be covered by disjoint occurrences of the $V_i)$.

Ergodically, the analog of the notion of rank with words replaced by patterns is called *funny rank*. Systems for which $T^* = 1$ are of *funny rank one*, some nontrivial example being given in [7]; systems for which $T^* > 0$ are of *local funny rank one*, and it is still true that they have simple spectrum if $T^* > \frac{1}{2}$.

We want now to compute T for the Morse sequence, so we have to look at the possible values of $p(V)\phi(V)$; for words, this quantity coincides with $p(V)f(V)$, and is bounded by $\frac{2}{3}$. For general patterns, the first thing we can do is to look at this $p(V)f(V)$, easier to compute than $p(V)\phi(V)$; this has the double interest of showing us that patterns do not behave like words, and of pointing to us which kind of patterns are more likely to be good tilers.

Proposition 12. *For every number a strictly smaller than $\frac{4}{3}$ and every integer n, we can find patterns U and V such that*

$$p(V)=2p(U)\geqslant n \quad \textit{and} \quad p(V)f(V)\geqslant ap(U)f(U).$$

The quantity $p(V)f(V)$ is unbounded on the set of finite patterns.

Proof. We extend to patterns the method we used to compute the frequency of words.

Let V be a pattern beginning by $u\overset{k}{\ldots}v\ldots$, suppose for example that k is even and nonzero; V occurs in some $\sigma^p e$ at place i if and only if

- either $i=2j$; then u is the end of a 1-word and v the beginning of a 1-word. So we know the pattern $u'u\overset{k}{\cdots}vv'\cdots$ occurs at place $i-1$ in $\sigma^p e$, and $u'\overset{k/2}{\ldots}v\ldots$ occurs at place j in $\sigma^{p-1}e$,
- or $i=2j-1$, and, in the same way, $u\overset{k/2-1}{\cdots}v\cdots$ must occur at place j in $\sigma^{p-1}e$.

Thus, we found two patterns V_1 and V_2 with

$$f(V)=\tfrac{1}{2}(f(V_1)+f(V_2)),$$

$$l(V_i)<l(V),$$

$$\tfrac{1}{2}p(V)\leqslant p(V_i)\leqslant p(V), \quad i=1,2.$$

we call them the *ancestors* of V; one of them may have frequency 0: we saw that a word of length bigger than 4 has only one ancestor. As the length of the ancestors is strictly smaller than the length of V, this process will ultimately give the frequency of V, as patterns of length l are words, and their frequency is known.

Now, let for any $k\geqslant 3$, $\tau(k)$ be the substitution on three symbols, 0, 1, and – (space), defined by

$$\tau(k)0=01\overset{2^k-2}{\cdots} \qquad \tau(k)1=10\overset{2^k-2}{\cdots} \qquad \tau(k)-=\overset{2^k}{\cdots}$$

(with the convention that we omit the final spaces of a pattern).

We shall apply the above method to compute $p(V_n)f(V_n)$ for $V_n=\tau(3)^n 0$: let $U=\tau(3)^n 0$ and $V=\tau(3)^{n+1}0$ for some $n\geqslant 1$; so U is of the form

$$U=uu'\overset{i}{\cdots}v\cdots$$

and so

$$V=uu'\overset{6}{\cdots}u'u\overset{8i+6}{\cdots}vv'\cdots$$

The previous method gives

$$f(V)=\tfrac{1}{2}f(u\overset{3}{\cdots}u'\overset{4i+3}{\cdots}v\cdots)+\tfrac{1}{2}f(u'u'\overset{2}{\cdots}uu\overset{4i+2}{\cdots}v'v'\cdots).$$

We have then

$$f(u'u'\overset{2}{\cdots}uu\overset{4i+2}{\cdots}v'v'\cdots)=\tfrac{1}{2}f(uu'u'u\overset{2i}{\cdots}vv'\cdots)=\tfrac{1}{4}f(uu'\overset{i}{\cdots}v\cdots)=\tfrac{1}{4}f(U),$$

the presence of $u'u'$ implying each time that there is only one ancestor;

$$f(u \overset{3}{\cdots} u' \overset{4i+3}{\cdots} v \cdots) = \tfrac{1}{2} f(u \overset{1}{\cdots} u' \overset{2i+1}{\cdots} v \cdots) + \tfrac{1}{2} f(u' \overset{1}{\cdots} u \overset{2i+1}{\cdots} v' \cdots)$$
$$= \tfrac{1}{4}(f(U) + f(U') + f(U') + f(U)) = f(U).$$

Hence, $f(V) = \frac{5}{8} f(U)$, and $p(V)f(V) = \frac{5}{4} p(U)f(U)$; hence $p(V_n)f(V_n)$ tends to infinity, which proves our second assertion.

To prove our first assertion we make the same computation with $\tau(3)$ replaced by $\tau(k)$; then $f(V) = \frac{1}{2} f(W) + \frac{1}{2} f(Z)$, where W is a pattern of same frequency as U, and Z has only one ancestor; but this unique ancestor of Z, which is of the form

$$uu' \overset{2^{k-2}-2}{\cdots} u'u \overset{2^{k-2}i+2^{k-2}-2}{\cdots} vv' \cdots$$

will in turn have two ancestors if k is strictly bigger than 3, giving one pattern of same frequency as U and one pattern looking like Z; finally, the expression giving $p(V)f(V)$ acquires a new term in $\frac{1}{16} p(U)f(U)$ when k is bigger than 6, a term in $\frac{1}{64} p(U)f(U)$ when k is bigger than 9, and so on. Hence, if k is big enough, $p(U)f(U) = ap(U)f(U)$, with a close to $(1 + \frac{1}{4} + \frac{1}{16} + \cdots) = \frac{4}{3}$. Hence our first assertion, as U and V are arbitrarily long if we chose n big enough. □

Corollary 13. *Lemma 7 is false for patterns.*

Proof. For patterns, $p(V)\phi(V) = t(V)$ is always smaller than one; hence if $p(V)f(V)$ is strictly bigger than one, the pattern V must have nondisjoint occurrences. □

For the privileged patterns we used, we are able to give estimates for the tiling:

Proposition 14. *There exists a sequence v_n, converging to $\frac{19}{24}$, such that*

$$t(\tau(3)^n 0) \geqslant v_n.$$

Proof. We compute the tiling frequency of the pattern $V_n = \tau(3)^n 0$.

V_1 being simply the word 01, has a tiling frequency of $\frac{2}{3}$, so there exists a subset A_1 of u, of measure $\frac{2}{3}$, covered by disjoint copies of V_1; also, there exists a subset B_1 of measure $\frac{2}{3}$ covered by disjoint copies of V_1', the word 10. Furthermore, $A_1 \cup B_1 = u$, because, if A_0 is the set of all 0's of u and B_0 the set of all 1's, we have $A_0 \cup B_0 = u$, $A_1 \supset \sigma A_0$, $B_1 \supset \sigma B_0$, so $u = \sigma u$ must be contained in $A_1 \cup B_1$.

At stage $n-1$, suppose we have found sets A_{n-1} and B_{n-1} satisfying

(1) $\mu(A_{n-1}) = \mu(B_{n-1}) = v_{n-1}$,

(2) $A_{n-1} \cup B_{n-1} = u$,

(3) A_{n-1} is tiled by V_{n-1} and B_{n-1} is tiled by V_{n-1}'.

Let M_{n-1} be the set of points x in u such that

(4) $x \in A_{n-1}$,

(5) $x \notin B_{n-1}$,

(6) the (unique) pattern V_{n-1} containing x in the tiling of A_{n-1} does not intersect B_{n-1}.

For a point x in u, $\sigma^3 x$ is a word of eight letters, namely $xx'x'xx'xxx'$, and we shall define A_n by the following rules:

(7) if x is in $A_{n-1} \cap B_{n-1}$, we put $\sigma^3 x$ in A_n,

(8) if x is in $B_{n-1} - A_{n-1}$, we put in A_n the underlined part of $\sigma^3 x$

$$xx'\underline{x'xx'x}xx',$$

(9) if x is in $A_{n-1} - B_{n-1} - M_{n-1}$, we put in A_n the underlined part of $\sigma^3 x$

$$\underline{xx'}x'xx'x\underline{xx'},$$

(10) if x is in M_{n-1}, we put in A_n the underlined part of $\sigma^3 x$

$$\underline{xx'}x'\underline{xx'}x\underline{xx'}.$$

We define B_n by exchanging, A_{n-1} and B_{n-1}, first in (4)–(6), defining M'_{n-1}, then in (7)–(10), replacing M_{n-1} by M'_{n-1}.

From the definition, we get that $A_n \cup B_n = u$, and that, if $w_{n-1} = \mu(M_{n-1})$

(11) $\mu(A_n) = v_n = v_{n-1} + \frac{1}{4} w_{n-1}$.

Also, we know V_{n-1} tiles A_{n-1}; furthermore, by (6), if a pattern V_{n-1} of this tiling is used to cover a part of M_{n-1}, this pattern must be wholly contained in M_{n-1}; in other words, V_{n-1} tiles M_{n-1}.

If x is a point in V_{n-1}, xx' is in V_n and if $x \overset{j}{\cdots} y$ is a subpattern of V_{n-1}, then $xx' \overset{8j+6}{\cdots} yy' = xx' \overset{3}{\cdots} \overset{8j}{\cdots} \overset{3}{\cdots} yy'$ is a subpattern of V_n; we have similar relations for x in V'_{n-1}. Therefore,

(12) as V_{n-1} tiles A_{n-1}, V_n tiles

$$(\underline{xx'}x'xx'x\underline{xx'},\ x \in A_{n-1}),$$

(13) as V_{n-1} tiles M_{n-1}, V_n tiles

$$(xx'x'\underline{xx'}xxx',\ x \in M_{n-1}),$$

(14) as V'_{n-1} tiles B_{n-1}, V_n tiles

$$(xx'\underline{x'xx'x}xx',\ x \in B_{n-1}).$$

These three sets cover A_n and are disjoint by construction; so we proved that V_n tiles A_n, and similarly that V'_n tiles B_n.

We need now to know the new M_n; we note that

$$A_n \cap B_n = (\sigma^3 x,\ x \in A_{n-1} \cap B_{n-1}) \cup (xx'x'\underline{xx'}xxx',\ x \in M_{n-1} \cap M'_{n-1}).$$

Now

- If x is in $A_{n-1} \cap B_{n-1}$, or in $A_{n-1} - B_{n-1} - M_{n-1}$, or in $B_{n-1} - A_{n-1} - M'_{n-1}$, x is covered by a pattern V_{n-1} (or V'_{n-1} in the last case) which intersects $A_{n-1} \cap B_{n-1}$.

So, by construction, the patterns V_n used to cover $\sigma^3 x$ intersects $A_n \cap B_n$; and so $\sigma^3 x$ cannot be in M_n.

- If x is in M'_{n-1}, then $xx'\underline{x'xx'x}xx'$ is in A_n, and $\underline{xx'}x'\underline{xx'}x\underline{xx'}$ is in B_n: only $xx'\underline{x'}xx'\underline{x}xx'$ is in $A_n - B_n$, but the pattern V_n used to cover that must contain $xx'\underline{x'xx'x}xx'$, hence $xx'x'\underline{xx'x}xx'$, which is in $A_n \cap B_n$; so $\sigma^3 x$ cannot be in M_n.
- If x is in M_{n-1}, then $\underline{xx'}x'xx'x\underline{xx'}$ is in $A_n - B_n$, and the patterns used to cover it do not intersect $A_n \cap B_n$, by definition of M_{n-1}.

Hence, we have proved that $\mu(M_n) = \frac{1}{2}\mu(M_{n-1})$.

Also we have the recurrence formulas, for $n \geqslant 1$,

$$V_{n+1} = v_n + \tfrac{1}{4} w_n,$$

$$w_{n+1} = \tfrac{1}{2} w_n;$$

at the beginning, $v_0 = \frac{1}{2}$ and $w_0 = \frac{1}{2}$, $A_0 \cap B_0$ being empty; then $v_1 = \frac{2}{3}$, and, by the same computation as for bigger n, we have $w_1 = \frac{1}{2} w_0$.

So $v_n = \frac{2}{3} + \frac{1}{4}\frac{1}{4} + \frac{1}{4}\frac{1}{8} + \cdots$,

$$v_n \to \tfrac{19}{24}, \quad \text{and} \quad v_n \leqslant t(V)_n. \quad \square$$

Proposition 15. *For the Morse sequence,* $T \geqslant \frac{5}{6}$.

Proof. We use the same method to compute $t(\tau(k)^n 0)$; in the same way, we can deduce the new tiled sets A_n and B_n from A_{n-1} and B_{n-1} by using the substitution σ^k; the set M_n has exactly the same definition, but the parts we can underline in (8)–(10) are slightly bigger in proportion, so that the recurrence relation is now

$$v_{n+1}(k) = v_n(k) + \alpha_k w_n(k),$$

$$w_{n+1}(k) = \tfrac{1}{2} w_n(k),$$

where α_k is *the proportion of the tiling of the k-words by the word* 01 *which does not come from a tiling of a* $k-1$*-word by the word* 0. (For $k=3$, the k-words are 01101001 and 10010110; if we want to tile these words by 01, we can cover $\underline{01}1010\underline{01}$ and $10\underline{0101}10$, but this tiling comes simply from applying the substitution to a tiling of the 2-words by the word 0; we can also cover $011\underline{01}001$, and this tiling does not come from a previous tiling by 0; the ratio between "new" and "previous" is $(\frac{1}{8})/(\frac{1}{2}) = \frac{1}{4} = \alpha_3$.)

When k tends to infinity, the tiling of the k-words by 01 tends to the tiling of the Morse sequence by 01, namely $\frac{2}{3}$; the part of this tiling which comes from a previous tiling by 0 tends to the tiling of the Morse sequence by 0, which is $\frac{1}{2}$; the ratio between new and previous tiling tends to $(\frac{2}{3} - \frac{1}{2})/(\frac{1}{2})$; so $\alpha_k \to \frac{1}{3}$.

So, when n and k go to infinity, the limit of the $v_n(k)$ is

$$\tfrac{2}{3} + \tfrac{1}{3}\cdot\tfrac{1}{4} + \tfrac{1}{3}\cdot\tfrac{1}{8} + \cdots = \tfrac{5}{6}$$

and so the upper limit of $t(\tau(k)^n 0)$ is bigger than $\frac{5}{6}$. $\square$

Of course, this gives only a minoration of T; however, for the particular patterns we used to tile, we can prove a more precise result; keeping the same notations as in the last proof, we have

Proposition 16. *$t(\tau(k)^n 0) = v_n(k)$ for all $k \geqslant 3$ and $n > 0$; in particular, the tilings of all these patterns are smaller than $\frac{5}{6}$.*

Proof. We write the proof only for $k=3$.

We begin by a precise study of $V_2 = 01 \overset{6}{\cdots} 10$; V_2 occurs in u in place i if and only if

$$i = 2k+1 \text{ and } 0 \overset{3}{\cdots} 1 \text{ occurs at place } k+1 \text{ or}$$

$$i = 2k \text{ and } 00 \overset{2}{\cdots} 11 \text{ occurs at place } k.$$

Or, continuing the process, we get that V_2 occurs at place i if and only if one of the following events takes place:

$$i = 8l+1 \text{ and } 01 \text{ occurs in place } l+1,$$

$$i = 8l+3 \text{ and } 10 \text{ occurs in place } l+1,$$

$$i = 8l+4 \text{ and } 01 \text{ occurs in place } l+1,$$

$$i = 8l+5 \text{ and } 10 \text{ occurs in place } l+1,$$

$$i = 8l+7 \text{ and } 01 \text{ occurs in place } l+1.$$

$\tau(3)$ being a substitution, this result is still true when we replace V_2 by V_{n+1}, 01 by V_n and 10 by V'_n, for any n bigger than 2.

Also so we see that any subset A of u tiled by V_n is deduced from a subset A' tiled by V_{n-1} and a subset B' tiled by V'_{n-1}, in the following way:

$$A \subset A_1 \cup B_1,$$

$$A_1 = (\underline{xx'}x'\underline{xx'}x\underline{xx'},\ x \in A'),$$

$$B_1 = (xx'\underline{x'xx'x}xx',\ x \in B').$$

$A_1 \cup B_1$ is covered by V_n but not tiled, as some elements of $A_1 \cap B_1$ may belong to two different patterns V_n; more precisely

$$A_1 \cap B_1 = (\sigma^3 x,\ x \in A' \cap B') \cup (xx'x'\underline{xx'x}xx',\ x \in A' \cup B')$$

and the elements of $A_1 \cup B_1$ belonging to two different patterns V_n (one coming from A' and one coming from B') form exactly the set

$$C_1 = (xx'x'\underline{xx'}xxx',\ x \in A' \cap B'),$$

A point in C_1 belongs to two different patterns V_n, α coming from A' and β coming from B'. To get the set A, we must delete α or β from the tiling, and thence delete from $A_1 \cup B_1$ also the part of α or β which is not in C_1.

However, for a point $xx'x'\underline{xx'}xxx'$ in C_1, to delete β would mean deleting all $xx'\underline{x'xx'x}xx'$ and so losing $\beta\cap\sigma^3(A'\cap B')$ *plus* a quarter of $\sigma^3 x$. To delete α means only losing $\alpha\cap\sigma^3(A'\cap B')$. Hence, the largest A we can build is obtained by deleting systematically the pattern α. Also in that case $\mu(A)\leqslant\mu(A'\cap B')+\frac{1}{2}\mu(B'-A')+\frac{1}{2}\mu(A'-B'-M')+\frac{3}{4}\mu(M')$, M' being defined as by (4)–(6).

This expression is optimal when $\mu(A')=\mu(B')$, and gives

$$\mu(A)\leqslant\mu(A')+\tfrac{1}{4}\mu(M').$$

Now A' is deduced in the same way from a set A'' tiled by V_{n-2}, involving a set M''. We could try, starting from A'' and M'', to build a *nonoptimal* A' in the hope of getting a bigger set M' and thus a bigger set A at the next stage; the previous analysis shows that the only possibility we have is, for some proportion of the x, to delete β instead of α (every other modification of the procedure would lessen $\mu(A')$ without increasing $\mu(M')$.

In this case there exists a number a in $[0,1]$ such that

$$\mu(A')\leqslant\mu(A'')+\tfrac{1}{4}\mu(M'')-a,$$

$$\mu(M')\leqslant\tfrac{1}{2}\mu(M'')+a,$$

which in turn gives

$$\mu(A)\leqslant\mu(A'')+\tfrac{1}{4}\mu(M'')-a+\tfrac{1}{4}(\tfrac{1}{2}\mu(M'')+a).$$

This value is optimal for $a=0$. So, as we know that $t(V_1)=v_1$, we get that at each stage $t(V_n)\leqslant v_n$. Hence, the proposition. □

Conjecture 17.

$$T=\tfrac{5}{6}.$$

Justification. We conjecture that the patterns $\tau(k)^n 0$ are the best tilers of the Morse sequence: to find a pattern which is a good tiler, the quantity $p(V)f(V)$, which is easy to compute, gives useful indications; $t(V)$ is *not* an increasing function of $p(V)f(V)$, for if $V=\tau(3)^n 0$ and $V'=\tau(100)^p 0$, for p big enough and n much bigger than p, we shall have $p(V)f(V)>p(V')f(V')$ but $t(V)<t(V')$ (this comes from Propositions 14 and 15 and the proof of Proposition 12), but from the proof of Propositions 14 and 15, it seems likely that the best tilers are sequences of patterns in which $p(V)f(V)$ grows very fast.

Now, if we try to compute the frequency of a given pattern V by the method we give, we see two extremal behaviours: if V is a word, it will have one ancestor, of weight (at best) $\frac{1}{2}p(V)$; if V is a "stretched" pattern, like $0\overset{3}{\cdots}1\overset{7}{\cdots}0$ (where each 0 or 1 is followed by at least one space), V will have two ancestors, both of weight $p(V)$. In the case of stretched patterns, there is no increase of $p(V)f(V)$ between ancestors and descendants, and hand computations with examples of manageable length seem to show that they do not tile better than words.

The interest lies in the opposite behaviours, that is patterns which behave "strictly better" than words; namely, patterns V which have two ancestors V_1 and V_2, with $p(V_1)=\frac{1}{2}p(V)$ and V_2 is of nonzero frequency; this is the case if and only if V is of the form

$$uu' \overset{4j+2}{\cdots} vv' \overset{4k+2}{\cdots} ww' \cdots$$

with $j>0$, $k>0, \cdots$

Now, V_2 is of the form

$$u'u' \overset{2j}{\cdots} v'v' \overset{2k}{\cdots} w'w' \cdots$$

and so $f(V_2)=\frac{1}{2}f(W)$, where W is

$$uu' \overset{j-1}{\cdots} vv' \overset{k-1}{\cdots} ww' \cdots$$

so W will be optimal when $j-1, k-1 \ldots$ are themselves of the form $4j'+2, 4k'+2, \ldots$ It is the iteration of that process which leads to the substitutions $\tau(k)$; hence our conjecture that they give the best possible tiling (it is also visible, in a more intuitive way, in the proof of Proposition 16: to improve a given tiling, we need to enlarge the picture, by applying some iterate σ^k of σ, and then try to fill the untiled space as well as we can; for fixed k, using $\tau(k)$ seems to be the best way of doing it).

For the computation of T^*, we should need some lemma analogous to Lemma 8 to show that a given pattern has little or no authorised neighbours; this is not clear as we use patterns whose length becomes big compared to their weight.

Problem 18. It would be interesting to compute F, F^*, T, T^* for other systems. Of course, we have the obvious inequalities

$$F \leqslant F^* \leqslant T^*, \qquad F \leqslant T \leqslant T^*,$$

on the other side, we know systems with $0<F^*<1$ [3], and systems with $T^*=1$ and $F^*=0$ [4].

For the Fibonacci substitution ($\sigma 0=01$, $\sigma 1=0$), we know that $F^*=T^*=1$. For the Rudin–Shapiro substitution ($\sigma 0=01$, $\sigma 1=02$, $\sigma 2=31$, $\sigma 3=32$), on the contrary, what we know of its ergodic properties, particularly the work of Quéffelec [11], implies that F, T, F^*, T^* must all be between $\frac{1}{4}$ and $\frac{1}{2}$.

A different open question is to know whether funny rank one is equivalent to simple spectrum. To find a system with $\frac{1}{2}<T^*<1$ would give a nice counterexample; as far as we know, this might even be the case for the Morse sequence itself.

References

[1] R.V. Chacon, Approximation and spectral multiplicity, in: *Contributions to Ergodic Theory and Probability*, Lecture Notes in Mathematics, Vol. 160 (Springer, Berlin, 1970) 18–27.

[2] A. Del Junco, A transformation with simple spectrum which is not rank one, *Canad. J. Math.* **29** (1977) 655–663.

[3] S. Ferenczi, Systèmes localement de rank un, *Ann. Inst. Henri Poincaré* **20** (1) (1984) 35–51.

[4] S. Ferenczi, Systèmes de rank un gauche, *Ann. Inst. Henri Poincaré* **21** (2) (1985) 177–186.

[5] G.A. Hedlund and M. Morse, Unending chess, symbolic dynamics and a problem in semi-groups, *Duke J. Math.* **11** (1944) 1–7.

[6] J.L. King, Joining-rank and the structure of finite-rank mixing transformations, *J. Analyse Math.* **51** (1988) 182–227.

[7] C. Mauduit, Substitutions et ensembles normaux, *Habilitation à diriger des recherches*, 1989, Université Aix-Marseille 2.

[8] M. Morse, Recurrent geodesics on a surface of negative curvature, *Trans. Amer. Math. Soc.* **22** (1921) 84–100.

[9] D.S. Ornstein, D.J. Rudolph and B. Weiss, Equivalence of measure-preserving transformations, *Memoirs Amer. Math. Soc.* **262** (1982).

[10] E. Prouhet, Mémoire sur quelques relations entre les puissances des nombres, *C.R. Acad. Sci. Paris* **33** (1851) 31.

[11] M. Queffelec, Substitution Dynamical Systems – Spectral Analysis, Lecture Notes in Mathematics, Vol. 1294 (Springer, Berlin, 1987).

[12] A. Thue, Über unendliche Zeichenreihen (1906), *Selected mathematical papers of Axel Thue*, Universitetsforlaget, 1977.